SEARCH REQUEST FORM

Requestor's Name: My-Chau Tran ⁷⁸⁹³³   Serial Number: 09/380,447

Date: 12/17/04   Phone: 272-0810   Art Unit: 1639

Mail box: REM C2A05   2C04

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

Mrs. O'Bryen,

Please seach claim 1 →

   and the following
   limitations:

1) The major coat protein is a M 13 filamentous phage (SEQ ID. No. 2)

2) Its variants are as follows:

~~delete~~

3) The heterologous polypeptide is an antibody that bind to a target, which is erb2.

4) The heterologous polypeptide is link to the major coat protein by a linking peptide of SEQ ID No. 110

                                    Thanks :))

S. Sidhu G. Weiss   J. Wells   Mod. of coat protein

STAFF USE ONLY

Date completed: 1-12-05

Searcher: BOB

Terminal time: seq 25 | txt 60

Elapsed time: 20 | 30

CPU time: _____

Total time: _____

Number of Searches: _____

Number of Databases: _____

Search Site

_____ STIC

_____ CM-1

_____ Pre-S

Type of Search

_____ N.A. Sequence

4 A.A. Sequence

_____ Structure

✗ Bibliographic

Vendors

_____ IG Suite

218 / 300 STN

_____ Dialog

_____ APS

_____ Geninfo

_____ SDC

_____ DARC/Questel

_____ Other

```
=> fil reg; d que 110
FILE 'REGISTRY' ENTERED AT 09:39:33 ON 12 JAN 2005
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2005 American Chemical Society (ACS)
```

Property values tagged with IC are from the ZIC/VINITI data file
provided by InfoChem.

```
STRUCTURE FILE UPDATES:   10 JAN 2005   HIGHEST RN 811411-12-8
DICTIONARY FILE UPDATES:  10 JAN 2005   HIGHEST RN 811411-12-8
```

TSCA INFORMATION NOW CURRENT THROUGH MAY 21, 2004

    Please note that search-term pricing does apply when
    conducting SmartSELECT searches.

Crossover limits have been increased. See HELP CROSSOVER for details.

Experimental and calculated property data are now available. For more
information enter HELP PROP at an arrow prompt in the file or refer
to the file summary sheet on the web at:
http://www.cas.org/ONLINE/DBSS/registryss.html

Seq 2, allowing for specified variations + conservative substitutions at the other positions

```
L10          13: SEA FILE=REGISTRY ABB=ON   [AD][EK][GS][DE][DK][PF][AS][KR][AD]A
             [FY][NE][SA]L[QE][AD][SI][AI]T[EN][YL][IF][GF][YL][AL][WG][AT][
             MV][VY]VVIVGATIGIKLFKKFTSKAS/SQSFP
```

```
=> d rn cn sql kwic nte lc 110 1-13
```

```
L10   ANSWER 1 OF 13  REGISTRY  COPYRIGHT 2005 ACS on STN
RN    799856-67-0  REGISTRY  - Use Registry # to match sequence to citation- (beginning on pg 6)
CN    7: PN: WO2004101608 SEQID: 7 unclaimed protein (9CI)   (CA INDEX NAME)
SQL   77
```

```
SEQ      1 SGTAMADPNR FRGKDLAGSP GGGSGGGAEG DDPAKAAFNS LQASATEYIG
                                           === ========== ==========

        51 YAWAMVVVIV GATIGIKLFK KFTSKAS
           ========== ========== =======
HITS AT:   28-77
```

RELATED SEQUENCES AVAILABLE WITH SEQLINK
```
LC    STN Files:   CA, CAPLUS, TOXCENTER
```

```
L10   ANSWER 2 OF 13  REGISTRY  COPYRIGHT 2005 ACS on STN
RN    685484-38-2  REGISTRY
CN    GenBank AAT27365 (9CI)   (CA INDEX NAME)        GenBank records printed at yellow tabs
OTHER NAMES:
CN    GenBank AAT27365 (Translated from: GenBank AY598820)
SQL   73
```

```
SEQ      1 MKKSLVLKAS VAVATLVPML SFAAEGDDPA KAAFNSLQAS ATEYIGYAWA
                                           ======= ========== ==========

        51 MVVVIVGATI GIKLFKKFTS KAS
           ========== ========== ===
HITS AT:   24-73
```

RELATED SEQUENCES AVAILABLE WITH SEQLINK

```
L10   ANSWER 3 OF 13  REGISTRY  COPYRIGHT 2005 ACS on STN
RN    480506-95-4  REGISTRY
```

CN   GenBank CAA23861 (9CI)   (CA INDEX NAME)
OTHER NAMES:
CN   GenBank CAA23861 (Translated from: GenBank V00604)
SQL   73

SEQ   1 MKKSLVLKAS VAVATLVPML SFAAEGDDPA KAAFNSLQAS ATEYIGYAWA
            ======= ========== ===========
         51 MVVVIVGATI GIKLFKKFTS KAS
            ========== ========== ===
HITS AT: 24-73

RELATED SEQUENCES AVAILABLE WITH SEQLINK

L10   ANSWER 4 OF 13   REGISTRY   COPYRIGHT 2005 ACS on STN
RN   460400-92-4   REGISTRY
CN   L-Serine, L-alanyl-L-α-glutamylglycyl-L-α-aspartyl-L-α-
      aspartyl-L-prolyl-L-alanyl-L-lysyl-L-alanyl-L-alanyl-L-phenylalanyl-L-
      asparaginyl-L-seryl-L-leucyl-L-glutaminyl-L-alanyl-L-seryl-L-alanyl-L-
      threonyl-L-α-glutamyl-L-tyrosyl-L-isoleucylglycyl-L-tyrosyl-L-alanyl-
      L-tryptophyl-L-alanyl-L-methionyl-L-valyl-L-valyl-L-valyl-L-isoleucyl-L-
      valylglycyl-L-alanyl-L-threonyl-L-isoleucylglycyl-L-isoleucyl-L-lysyl-L-
      leucyl-L-phenylalanyl-L-lysyl-L-lysyl-L-phenylalanyl-L-threonyl-L-seryl-L-
      lysyl-L-alanyl- (9CI)   (CA INDEX NAME)
OTHER NAMES:
CN   3: PN: WO02072780 SEQID: 10 unclaimed protein
SQL   50

SEQ   1 AEGDDPAKAA FNSLQASATE YIGYAWAMVV VIVGATIGIK LFKKFTSKAS
            ========== ========== ========== ========== ==========
HITS AT: 1-50

RELATED SEQUENCES AVAILABLE WITH SEQLINK
LC   STN Files: CA, CAPLUS, TOXCENTER, USPATFULL

L10   ANSWER 5 OF 13   REGISTRY   COPYRIGHT 2005 ACS on STN
RN   405053-51-2   REGISTRY
CN   L-Serine, L-alanyl-L-α-glutamylglycyl-L-α-aspartyl-L-α-
      aspartyl-L-prolyl-L-alanyl-L-lysyl-L-alanyl-L-alanyl-L-phenylalanyl-L-
      asparaginyl-L-seryl-L-leucyl-L-glutaminyl-L-alanyl-L-seryl-L-alanyl-L-
      threonyl-L-α-glutamyl-L-tyrosyl-L-isoleucylglycyl-L-tyrosyl-L-alanyl-
      L-tryptophyl-L-alanyl-L-methionyl-L-valyl-L-valyl-L-valyl-L-isoleucyl-L-
      valylglycyl-L-alanyl-L-threonyl-L-isoleucylglycyl-L-isoleucyl-L-lysyl-L-
      leucyl-L-phenylalanyl-L-lysyl-L-lysyl-L-phenylalanyl-L-threonyl-L-seryl-L-
      lysyl-L-alanyl- (9CI)   (CA INDEX NAME)
SQL   50

SEQ   1 AEGDDPAKAA FNSLQASATE YIGYAWAMVV VIVGATIGIK LFKKFTSKAS
            ========== ========== ========== ========== ==========
HITS AT: 1-50

RELATED SEQUENCES AVAILABLE WITH SEQLINK
LC   STN Files: CA, CAPLUS

L10   ANSWER 6 OF 13   REGISTRY   COPYRIGHT 2005 ACS on STN
RN   346502-45-2   REGISTRY
CN   21: PN: US6251865 SEQID: 21 unclaimed protein (9CI)   (CA INDEX NAME)
OTHER NAMES:
CN   1: PN: WO0187323 PAGE: 39 unclaimed sequence
SQL   77

SEQ   1 SGTAMADPNR FRGKDLAGSP GGGSGGGAEG DDPAKAAFNS LQASATEYIG
                        === ========== ==========
         51 YAWAMVVVIV GATIGIKLFK KFTSKAS

Searched by Barb O'Bryen, STIC 2-2518

```
========== ========== =======
```
HITS AT: 28-77

RELATED SEQUENCES AVAILABLE WITH SEQLINK
LC   STN Files: CA, CAPLUS, TOXCENTER, USPATFULL

L10   ANSWER 7 OF 13   REGISTRY   COPYRIGHT 2005 ACS on STN
RN   316203-10-8   REGISTRY
CN   Protein A (synthetic Staphylococcus aureus clone pAlg-816 precursor
      N-terminal fragment) fusion protein with immunoglobulin, anti-(Z-DNA)
      (mouse single-chain) fusion protein with protein (coliphage M13 gene III
      coat C-terminal fragment) (9CI)   (CA INDEX NAME)
SQL   364

SEQ   301 KLEIKHHHHH GSGTAEGDDP AKAAFNSLQA SATEYIGYAW AMVVVIVGAT
               ====== ========== ========== ==========

       351 IGIKLFKKFT SKAS
           ========== ====
HITS AT: 315-364
LC   STN Files: CA, CAPLUS

L10   ANSWER 8 OF 13   REGISTRY   COPYRIGHT 2005 ACS on STN
RN   151033-87-3   REGISTRY
CN   (1-6)-(9-21)-Phosphatase, alkaline (Escherichia coli precursor reduced)
      fusion protein with trypsin inhibitor (cattle pancreas basic reduced)
      fusion protein with protein (coliphage M13 coat) (9CI)   (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN   (1-6)-(9-21)-Phosphatase, alkaline (Escherichia coli precursor reduced)
      fusion protein with trypsin inhibitor (ox pancreas basic reduced) fusion
      protein with protein (coliphage M13 coat)
OTHER NAMES:
CN   (1-6)-(9-21)-Phosphatase, alkaline (Escherichia coli precursor reduced)
      fusion protein with trypsin inhibitor (ox pancreas basic reduced) fusion
      protein with protein (bacteriophage M13 coat)
SQL   127

SEQ   51 TFVYGGCRAK RNNFKSAEDC MRTCGGAAEG DDPAKAAFNS LQASATEYIG
                                   === ========== ==========

       101 YAWAMVVVIV GATIGIKLFK KFTSKAS
           ========== ========== =======
HITS AT: 78-127
LC   STN Files: CA, CAPLUS, TOXCENTER, USPATFULL

L10   ANSWER 9 OF 13   REGISTRY   COPYRIGHT 2005 ACS on STN
RN   144999-51-9   REGISTRY
CN   Trypsin inhibitor (cattle pancreas basic reduced), (58→1')-protein
      with protein (coliphage M13 coat) (9CI)   (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN   Trypsin inhibitor (ox pancreas basic reduced), (58→1')-protein with
      protein (coliphage M13 coat)
OTHER NAMES:
CN   Trypsin inhibitor (ox pancreas basic reduced), (58→1')-protein with
      protein (bacteriophage M13 coat)
SQL   108

SEQ   51 CMRTCGGAAE GDDPAKAAFN SLQASATEYI GYAWAMVVVI VGATIGIKLF
               == ========== ========== ========== ==========

       101 KKFTSKAS
           ========
HITS AT: 59-108
LC   STN Files: CA, CAPLUS, TOXCENTER

L10   ANSWER 10 OF 13   REGISTRY   COPYRIGHT 2005 ACS on STN

RN   144999-50-8 REGISTRY
CN   1-21-Phosphatase, alkaline (Escherichia coli precursor reduced),
     (21→1')-protein with trypsin inhibitor (cattle pancreas basic
     reduced) (58'→1'')-protein with protein (coliphage M13 coat) (9CI)
     (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN   1-21-Phosphatase, alkaline (Escherichia coli precursor reduced),
     (21→1')-protein with trypsin inhibitor (ox pancreas basic reduced)
     (58'→1'')-protein with protein (coliphage M13 coat)
OTHER NAMES:
CN   1-21-Phosphatase, alkaline (Escherichia coli precursor reduced),
     (21→1')-protein with trypsin inhibitor (ox pancreas basic reduced)
     (58'→1'')-protein with protein (bacteriophage M13 coat)
SQL 129

SEQ   51 CQTFVYGGCR AKRNNFKSAE DCMRTCGGAA EGDDPAKAAF NSLQASATEY
            = ========== ==========
        101 IGYAWAMVVV IVGATIGIKL.FKKFTSKAS
            ========== ========== =========
HITS AT: 80-129
LC   STN Files: CA, CAPLUS, TOXCENTER

L10 ANSWER 11 OF 13 REGISTRY COPYRIGHT 2005 ACS on STN
RN   114013-90-0 REGISTRY
CN   1-142-Kinase (phosphorylating), ribulo- (Salmonella typhimurium reduced)
     142-glycine-, (142→11')-protein with 11-73-protein (coliphage M13
     gene 8) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN   1-142-Kinase (phosphorylating), ribulo- (Salmonella typhimurium reduced)
     142-glycine-, (142→11')-protein with 11-73-protein (bacteriophage
     M13 gene 8)
SQL 205

SEQ   151 MLSFAAEGDD PAKAAFNSLQ ASATEYIGYA WAMVVVIVGA TIGIKLFKKF
            ===== ========== ========== ========== ==========
        201 TSKAS
            =====
HITS AT: 156-205
LC   STN Files: CA, CAPLUS

L10 ANSWER 12 OF 13 REGISTRY COPYRIGHT 2005 ACS on STN
RN   104364-70-7 REGISTRY
CN   Protein (coliphage M13 coat) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN   2: PN: WO0006717 SEQID: 2 unclaimed protein
CN   Protein (bacteriophage M13 coat)
SQL 50

SEQ   1 AEGDDPAKAA FNSLQASATE YIGYAWAMVV VIVGATIGIK LFKKFTSKAS
            ========== ========== ========== ========== ==========
HITS AT: 1-50

RELATED SEQUENCES AVAILABLE WITH SEQLINK
LC   STN Files: CA, CAPLUS, TOXCENTER

L10 ANSWER 13 OF 13 REGISTRY COPYRIGHT 2005 ACS on STN
RN   76560-64-0 REGISTRY
CN   Protein (coliphage M13 coat precursor) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN   Protein (bacteriophage M13 coat precursor)
SQL 73

SEQ   1 MKKSLVLKAS VAVATLVPML SFAAEGDDPA KAAFNSLQAS ATEYIGYAWA

```
                    ======= ========== ==========
       51 MVVVIVGATI GIKLFKKFTS KAS
          ========== ========== ===
HITS AT:   24-73
```

RELATED SEQUENCES AVAILABLE WITH SEQLINK
LC   STN Files: CA, CAPLUS

=> => fil capl; s l10
FILE 'CAPLUS' ENTERED AT 09:40:26 ON 12 JAN 2005
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2005 AMERICAN CHEMICAL SOCIETY (ACS)

Registry file answer set crossed into bibliographic files to get citations

Copyright of the articles to which records in this database refer is
held by the publishers listed in the PUBLISHER (PB) field (available
for records published or updated in Chemical Abstracts after December
26, 1996), unless otherwise indicated in the original publications.
The CA Lexicon is the copyrighted intellectual property of the
American Chemical Society and is provided to assist you in searching
databases on STN. Any dissemination, distribution, copying, or storing
of this information, without the prior written consent of CAS, is
strictly prohibited.

FILE COVERS 1907 - 12 Jan 2005 VOL 142 ISS 3
FILE LAST UPDATED: 10 Jan 2005 (20050110/ED)

   This file contains CAS Registry Numbers for easy and accurate
   substance identification.

'OBI' IS DEFAULT SEARCH FIELD FOR 'CAPLUS' FILE

L15   12 L10

=> fil uspatf toxcenter; s l10
FILE 'USPATFULL' ENTERED AT 09:40:36 ON 12 JAN 2005
CA INDEXING COPYRIGHT (C) 2005 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'TOXCENTER' ENTERED AT 09:40:36 ON 12 JAN 2005
COPYRIGHT (C) 2005 ACS

L16   9 L10

=> dup rem l15,l16
FILE 'CAPLUS' ENTERED AT 09:40:40 ON 12 JAN 2005
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2005 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'USPATFULL' ENTERED AT 09:40:40 ON 12 JAN 2005
CA INDEXING COPYRIGHT (C) 2005 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'TOXCENTER' ENTERED AT 09:40:40 ON 12 JAN 2005
COPYRIGHT (C) 2005 ACS
PROCESSING COMPLETED FOR L15
PROCESSING COMPLETED FOR L16
L17   13 DUP REM L15 L16 (8 DUPLICATES REMOVED)
                ANSWERS '1-12' FROM FILE CAPLUS
                ANSWER '13' FROM FILE USPATFULL

                 Searched by Barb O'Bryen, STIC 2-2518
```

OM protein - protein search, using sw model

Copyright (c) 1993 - 2005 Compugen Ltd.
GenCore version 5.1.6

Run on: January 12, 2005, 10:08:53

Title:          SEQ110-THEN-SEQ2
Perfect score:  304
Sequence:       1 AWEENIDSAPAEGDDPAKAA.........VIVGATIGIKLFKKFTSKAS 60

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.1
Searched:       2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters:        2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      A_Geneseq_23Sep04:*
                1: geneseqp1980s:*
                2: geneseqp1990s:*
                3: geneseqp2000s:*
                4: geneseqp2001s:*
                5: geneseqp2002s:*
                6: geneseqp2003as:*
                7: geneseqp2003bs:*
                8: geneseqp2004s:*

Search time 115 Seconds
(without alignments)
187.163 Million cell updates/sec

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 250.1 | 82.3 | 131 | 2 | AAR04004 | Aar04004 Portion o |
| 2 | 249 | 81.9 | 73 | 2 | AAR04003 | Aar04003 Portion o |
| 3 | 248 | 81.6 | 50 | 2 | AAR50175 | Aar50175 Membrane |
| 4 | 248 | 81.6 | 50 | 2 | ABJ15250 | Abj15250 Peptide o |
| 5 | 248 | 81.6 | 77 | 2 | AAW67493 | Aaw67493 gD-tag fo |
| 6 | 248 | 81.6 | 77 | 4 | AAG63727 | Aag63727 Protein o |
| 7 | 248 | 81.6 | 77 | 5 | AAM48216 | Aam48216 Antibody- |
| 8 | 248 | 81.6 | 77 | 5 | ADD84797 | Add84797 Polypepti |
| 9 | 248 | 81.6 | 77 | 7 | ADH48110 | Adh48110 Antibody |
| 10 | 248 | 81.6 | 77 | 8 | ADF47603 | Adf47603 Insulin-1 |
| 11 | 248 | 81.6 | 77 | 8 | ADG39549 | Adg39549 Phage dis |
| 12 | 248 | 81.6 | 77 | 8 | ADI62030 | Adi62030 Antibody- |
| 13 | 248 | 81.6 | 77 | 8 | ADI57543 | Adi57543 Plasmid p |
| 14 | 248 | 81.6 | 77 | 8 | ADI57436 | Adi57436 Plasmid p |
| 15 | 248 | 81.6 | 77 | 8 | ADJ56592 | Adj56592 Plasmid p |
| 16 | 248 | 81.6 | 77 | 8 | ADJ58114 | Adj58114 Random-se |
| 17 | 248 | 81.6 | 77 | 8 | ADJ45114 | Adj45114 Plasmid p |
| 18 | 248 | 81.6 | 77 | 8 | ADM76458 | Adm76458 Phage dis |
| 19 | 248 | 81.6 | 77 | 6 | ADL08228 | Adl08228 Plasmid p |
| 20 | 242.7 | 79.8 | 71 | 6 | ABU09979 | Abu09979 Phagemid |
| 21 | 234 | 77.0 | 72 | 3 | AAY94389 | Aay94389 Fusion pr |
| 22 | 233.9 | 76.9 | 49 | 2 | AAR28600 | Aar28600 Sequence |
| 23 | 233 | 76.6 | 132 | 8 | AAY81275 | Aay81275 Mature M1 |
| 24 | 229 | 75.3 | 73 | 1 | ADD14860 | Add14860 Dual disp |
| 25 | 226.1 | 74.4 | 79 | 8 | ADP64647 | Adp64647 HIV-1 sub |
| 26 | 93 | 30.6 | 20 | 7 | ADM34212 | Adm34212 Synthetic |
| 27 | 88 | 28.9 | 24 | 7 | ABR83076 | Abr83076 Major coa |
| 28 | 71.4 | 23.5 | 25 | 7 | ABP65108 | Abp65108 Hypoxia-i |
| 29 | 71.4 | 23.5 | 625 | 7 | ADD14192 | Add14192 Human src |
| 30 | 71.4 | 23.5 | 625 | 8 | ADN04317 | Adn04317 Antipsori |
| 31 | 71.4 | 23.5 | 653 | 7 | ADJ70738 | Adj70738 Human hea |
| 32 | 71.4 | 23.5 | 677 | 7 | ADE09955 | Ade09955 Novel pro |
| 33 | 69.1 | 22.7 | 491 | 4 | AAG91147 | Aag91147 C_glutami |
| 34 | 69.1 | 22.7 | 533 | 7 | ADL65983 | Adl65983 Novel pro |
| 35 | 67.9 | 22.3 | 2146 | 4 | ABB62317 | Abb62317 Drosophil |
| 36 | 67.8 | 22.3 | 618 | 4 | ABG09795 | Abg09795 Novel hum |
| 37 | 67.6 | 22.2 | 303 | 7 | ADH86191 | Adh86191 Enterococ |
| 38 | 66.9 | 22.0 | 476 | 5 | ABB91267 | Abb91267 Herbicida |
| 39 | 66.9 | 22.0 | 476 | 8 | ADJ49077 | Adj49077 Oil-assoc |
| 40 | 66.9 | 22.0 | 480 | 8 | ADJ49984 | Adj49984 Oil-assoc |
| 41 | 66.9 | 22.0 | 480 | 8 | ADJ50237 | Adj50237 Oil-assoc |
| 42 | 66.3 | 21.8 | 480 | 3 | AAG18633 | Aag18633 Arabidops |
| 43 | 66.3 | 21.8 | 464 | 3 | AAG46017 | Aag46017 Arabidops |
| 44 | 66.3 | 21.8 | 464 | 3 | AAG18632 | Aag18632 Arabidops |
| 45 | 66.3 | 21.8 | 496 | 3 | AAG46016 | Aag46016 Arabidops |

ALIGNMENTS

RESULT 1
AAR04004

ID   AAR04004 standard; protein; 131 AA.
AC   AAR04004;
XX
XX
DT   30-AUG-1990   (first entry)
XX
DE   Portion of M13 coat protein with bovine pancreatic trypsin inhibitor
DE   insert.
XX
KW   Binding proteins; M13cp-BPTI fusion gene; horse heart myoglobin;
KW   Potential binding domain.
XX
OS   Synthetic.
XX
FH   Key              Location/Qualifiers
FT   Region           24..81
FT                    /label= BPTI sequence
XX
XX
PN   WO900 2809-A.
XX
PD   22-MAR-1990.
XX
PF   02-SEP-1988;  88US-00240160.
XX
PR   02-SEP-1988;  88US-00240160.
XX
PA   (PROT-) PROTEIN ENG CORP.
XX
PI   Ladner RC, Guterman SK;
XX
DR   WPI; 1990-115996/15.
XX
PT   Prepn. of proteins which bind predetermined target - by generation and
PT   selection of recombinant varied binding proteins using replicable genetic
PT   packages.
XX
PS   Disclosure; Page 133; 269pp; English.
XX
CC   An insert is made into the M13 coat protein whereby the phage will
CC   present a binding domain as a surface protein. That sequence is
CC   specifically of the bovine pancreatic trypsin inhibitor (BPTI) which has
CC   affinity for an horse heart myoglobin target protein. See also AAQ03954-
58
SQ   Sequence 131 AA;

_[Handwritten marginal annotations are present in the center of the page, including "Seq 110", "Seq 2", "Seq 110 & Seq 2", and further illegible notes.]_

OM protein - protein search, using sw model
                    GenCore version 5.1.6
                    Copyright (c) 1993 - 2005 Compugen Ltd.

Run on:             January 12, 2005, 10:43:39 ; Search time 28 Seconds
                                                (without alignments)
                                                142.110 Million cell updates/sec

Title:              SEQ110-THEN-SEQ2
Perfect score:      304
Sequence:           1 AWEENIDSAPAEGDDPAKAA..........VIVGATIGIKLFKKFTSKAS 60

Scoring table:      BLOSUM62
                    Gapop 10.0 , Gapext 0.1

Searched:           478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters:     478139

Post-processing: Minimum Match 0%
Maximum Match 100%

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Database :          Issued_Patents_AA:*
                    1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
                    2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
                    3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
                    4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
                    5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
                    6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

                    Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Match Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 248 | 81.6 | 50 | 1 | US-08-133-011-17 | Sequence 17, Appl |
| 2 | 248 | 81.6 | 50 | 1 | US-08-322-730A-17 | Sequence 17, Appl |
| 3 | 248 | 81.6 | 50 | 1 | US-08-387-874-17 | Sequence 17, Appl |
| 4 | 248 | 81.6 | 50 | 2 | US-08-383-619-17 | Sequence 17, Appl |
| 5 | 248 | 81.6 | 50 | 3 | US-08-907-739-17 | Sequence 17, Appl |
| 6 | 248 | 81.6 | 50 | 5 | US-09-729-597-17 | Sequence 17, Appl |
| 7 | 248 | 81.6 | 50 | 5 | PCT-US93-08364-17 | Sequence 17, Appl |
| 8 | 248 | 81.6 | 77 | 3 | US-08-825-852-21 | Sequence 21, Appl |
| 9 | 248 | 81.6 | 77 | 4 | US-09-052-888-21 | Sequence 21, Appl |
| 10 | 248 | 81.6 | 77 | 4 | US-09-723-890-21 | Sequence 21, Appl |
| 11 | 248 | 81.6 | 77 | 4 | US-09-723-901-21 | Sequence 21, Appl |
| 12 | 248 | 81.6 | 77 | 4 | US-09-723-547-21 | Sequence 21, Appl |
| 13 | 248 | 81.6 | 77 | 4 | US-09-724-127-21 | Sequence 21, Appl |
| 14 | 248 | 81.6 | 77 | 4 | US-09-723-931-21 | Sequence 21, Appl |
| 15 | 248 | 81.6 | 77 | 4 | US-09-723-873-21 | Sequence 21, Appl |
| 16 | 248 | 81.6 | 77 | 4 | US-09-724-114-21 | Sequence 21, Appl |
| 17 | 248 | 81.6 | 77 | 4 | US-09-723-913-21 | Sequence 21, Appl |
| 18 | 248 | 81.6 | 77 | 4 | US-09-723-912-21 | Sequence 21, Appl |
| 19 | 248 | 81.6 | 77 | 4 | US-09-724-095-21 | Sequence 21, Appl |
| 20 | 248 | 81.6 | 77 | 4 | US-09-724-157-21 | Sequence 21, Appl |
| 21 | 248 | 81.6 | 77 | 4 | US-09-724-062-21 | Sequence 21, Appl |
| 22 | 248 | 81.6 | 77 | 4 | US-09-724-065-21 | Sequence 21, Appl |
| 23 | 244 | 80.3 | 73 | 4 | US-09-495-880A-9 | Sequence 9, Appl |
| 24 | 244 | 80.3 | 73 | 4 | US-09-495-880A-24 | Sequence 24, Appl |
| 25 | 244 | 80.3 | 73 | 4 | US-09-495-880A-40 | Sequence 40, Appl |
| 26 | 234 | 77.0 | 72 | 4 | US-09-866-073A-17 | Sequence 17, Appl |
| 27 | 71.4 | 23.5 | 625 | 4 | US-09-538-092-1064 | Sequence 1064, Ap |

---

OM protein - protein search, using sw model
                    GenCore version 5.1.6
                    Copyright (c) 1993 - 2005 Compugen Ltd.

Run on:             January 12, 2005, 11:00:20 ; Search time 97.5 Seconds
                                                (without alignments)
                                                222.332 Million cell updates/sec

Title:              SEQ110-THEN-SEQ2
Perfect score:      304
Sequence:           1 AWEENIDSAPAEGDDPAKAA..........VIVGATIGIKLFKKFTSKAS 60

Scoring table:      BLOSUM62
                    Gapop 10.0 , Gapext 0.1

Searched:           1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters:     1608061

Post-processing: Minimum Match 0%
Maximum Match 100%

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database :          Published_Applications_AA:*
                    1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
                    2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
                    3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
                    4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
                    5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
                    6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
                    7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
                    8: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
                    9: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
                    10: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
                    11: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
                    12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
                    13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
                    14: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
                    15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
                    16: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
                    17: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
                    18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
                    19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
                    20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

                    Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Match Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 255 | 83.9 | 80 | 16 | US-10-416-708A-67 | Sequence 67, Appl |
| 2 | 250.1 | 82.3 | 127 | 10 | US-09-896-095-270 | Sequence 270, App |
| 3 | 250.1 | 82.3 | 131 | 10 | US-09-896-095-187 | Sequence 187, App |
| 4 | 250.1 | 82.3 | 131 | 10 | US-09-896-095-268 | Sequence 268, App |
| 5 | 250.1 | 82.3 | 131 | 10 | US-09-896-095-273 | Sequence 273, App |
| 6 | 250.1 | 82.3 | 132 | 10 | US-09-896-095-220 | Sequence 220, App |
| 7 | 250.1 | 82.3 | 132 | 10 | US-09-896-095-222 | Sequence 222, App |
| 8 | 250.1 | 82.3 | 132 | 10 | US-09-896-095-224 | Sequence 224, App |
| 9 | 249 | 81.9 | 73 | 10 | US-09-896-095-122 | Sequence 122, App |
| 10 | 249 | 81.9 | 73 | 10 | US-10-098-093-10 | Sequence 10, Appl |
| 11 | 248 | 81.6 | 50 | 14 | US-10-273-973-17 | Sequence 17, Appl |
| 12 | 248 | 81.6 | 77 | 10 | US-09-858-935B-2 | Sequence 2, Appli |
| 13 | 248 | 81.6 | 77 | 10 | US-09-858-935B-55 | Sequence 55, Appl |

## Left report

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: January 12, 2005, 10:11:18 ; Search time 126 Seconds
(without alignments)
273.988 Million cell updates/sec

Title:          SEQ110-THEN-SEQ2
Perfect score:  304
Sequence:       1 AWEENIDSAPAEGDDPAKAA.........VIVGATIGIKLFKKFPTSKAS 60

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.1

Searched:       1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters:      1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      Uniprot_02:*
                1:  uniprot_sprot:*
                2:  uniprot_trembl:*,

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Match Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 244 | 80.3 | 73 | 1 | COAB_BPFD | P03617 bacteriopha |
| 2 | 243 | 79.9 | 50 | 2 | Q9T0Q9 | Q9T0Q9 bacteriopha |
| 3 | 238 | 78.3 | 50 | 1 | COAB_BP2J2 | P03618 bacteriopha |
| 4 | 206 | 67.8 | 65 | 1 | COAB_BPIF1 | Q38211 bacteriopha |
| 5 | 173.7 | 57.1 | 74 | 1 | COAB_BPIF1 | P03619 bacteriopha |
| 6 | 97 | 31.9 | 84 | 1 | COAB_BP122 | P15416 bacteriopha |
| 7 | 93 | 30.6 | 53 | 2 | Q9T0Q8 | Q9T0Q8 bacteriopha |
| 8 | 90 | 29.6 | 82 | 1 | COAB_BPIKE | P03620 bacteriopha |
| 9 | 90 | 29.6 | 82 | 1 | COAB_BPIKE | P03620 bacteriopha |
| 10 | 75.2 | 24.7 | 183 | 2 | Q9F4D4 | Q9F4D4 burkholderi |
| 11 | 75.1 | 24.7 | 1070 | 2 | Q6I5X9 | Q6I5X9 oryza sativ |
| 12 | 73.5 | 24.2 | 476 | 1 | Q9A3A6 | Q9A3A6 caulobacter |
| 13 | 72.8 | 23.9 | 351 | 1 | O9H187 | O9h187 thermoplasm |
| 14 | 71.7 | 23.6 | 535 | 1 | YE55_THERAC | O55577 synechocyst |
| 15 | 71.4 | 23.5 | 252 | 2 | O8C514 | O8C514 mus musculu |
| 16 | 71.4 | 23.5 | 504 | 2 | O60786 | O60786 mus musculu |
| 17 | 71.4 | 23.5 | 540 | 2 | O81S7 | O81S7 mus musculu |
| 18 | 71.4 | 23.5 | 653 | 1 | M1A1_HUMAN | P33908 homo sapien |
| 19 | 71.1 | 23.5 | 1031 | 2 | O8EFT4 | O8eft4 shewanella |
| 20 | 71.1 | 23.4 | 228 | 2 | O73U38 | O73U38 mycobacteri |
| 21 | 70.7 | 23.4 | 228 | 2 | AAS06080 | AAS06080 mycobacteri |
| 22 | 70.7 | 23.3 | 2167 | 2 | O92EK5 | O92EK5 listeria in |
| 23 | 70.3 | 23.1 | 228 | 2 | O89114 | O89114 bradyrhizob |
| 24 | 70.2 | 23.1 | 400 | 2 | O892A4 | O892a4 bradyrhizob |
| 25 | 70.2 | 23.1 | 436 | 2 | O73P16 | O73P16 bacteroides |
| 26 | 69.9 | 23.0 | 499 | 2 | AAS11474 | AAS11474 treponema d |
| 27 | 69.5 | 22.9 | 499 | 2 | O7NSA1 | O7NSA1 treponema |
| 28 | 69.4 | 22.8 | 700 | 2 | O89ED0 | O89ED0 chromobacte |
| 29 | 69.4 | 22.8 | 236 | 2 | O7VBS1 | O7VBS1 bradyrhizob |
| 30 | 69.4 | 22.8 | 523 | 2 | O6Y132 | O6Y132 prochloroco |
| 31 | 69.3 | 22.8 | 1285 | 2 | AAQ72578 | AAQ72578 lactuca sat |
|  |  |  | 1285 | 2 | Q6C6N0 | Q6C6N0 yarrowia li |

## Right report

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: January 12, 2005, 10:42:14 ; Search time 24.5 Seconds
(without alignments)
235.633 Million cell updates/sec

Title:          SEQ110-THEN-SEQ2
Perfect score:  304
Sequence:       1 AWEENIDSAPAEGDDPAKAA.........VIVGATIGIKLFKKFPTSKAS 60

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.1

Searched:       283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters:      283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      PIR_79:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Match Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 249 | 81.9 | 73 | 1 | VCBPM3 | coat protein B pre |
| 2 | 244 | 80.3 | 73 | 1 | VCBPF1 | coat protein B pre |
| 3 | 244 | 80.3 | 73 | 1 | VCBPFD | coat protein B pre |
| 4 | 238 | 78.3 | 50 | 1 | VCBPZ2 | coat protein B pre |
| 5 | 170.6 | 56.1 | 51 | 1 | VCBPIF | coat protein - pha |
| 6 | 97 | 31.9 | 84 | 2 | VCBPIK | coat protein - p |
| 7 | 90 | 29.6 | 82 | 2 | S08090 | gene VIII protein |
| 8 | 90 | 29.6 | 82 | 2 | VCBPIK | coat protein B pre |
| 9 | 73.5 | 24.2 | 476 | 1 | H87657 | conserved hypothet |
| 10 | 71.7 | 23.6 | 535 | 1 | S76103 | hypothetical prote |
| 11 | 71.4 | 23.5 | 504 | 2 | I49257 | hypothetical prote |
| 12 | 71.4 | 23.3 | 625 | 2 | S38965 | NF2d9 -- mouse |
| 13 | 70.7 | 23.0 | 419 | 2 | AFI489 | mannosyl-oligosacc |
| 14 | 69.1 | 22.4 | 2167 | 2 | TI9260 | cell wall-associat |
| 15 | 68 | 22.4 | 513 | 2 | TI5899 | hypothetical prote |
| 16 | 67.9 | 22.3 | 520 | 2 | T43686 | cytochrome-c oxida |
| 17 | 67.6 | 22.2 | 2139 | 2 | A35672 | crumbs protein - f |
| 18 | 66.4 | 22.0 | 446 | 2 | C26505 | NADH oxidase - Ent |
| 19 | 66 | 21.8 | 476 | 2 | S26965 | probable amino aci |
| 20 | 66 | 21.7 | 505 | 2 | D70799 | hypothetical prote |
| 21 | 65.7 | 21.6 | 369 | 2 | F87634 | dihydroorotate oxi |
| 22 | 65.7 | 21.6 | 317 | 2 | A33559 | conserved hypothet |
| 23 | 65.6 | 21.6 | 411 | 2 | C57479 | amino acid transpo |
| 24 | 65.6 | 21.6 | 598 | 2 | T45081 | probable transmemb |
| 25 | 65.2 | 21.4 | 388 | 2 | TF0288 | 3-ketoacyl-CoA thi |
| 26 | 65 | 21.4 | 5149 | 2 | F83345 | probable non-ribos |
| 27 | 65 | 21.4 | 400 | 2 | C85882 | permease of transp |
| 28 | 64.9 | 21.3 | 910 | 2 | H91037 | permease of transp |
| 29 | 64.9 | 21.3 | 1331 | 2 | JC4609 | chitin synthase (B |
|  | 64.9 | 21.3 | 1331 | 2 | T18310 | receptor-adenylate |